

GÖTEBORGS UNIVERSITET
PSYKOLOGISKA INSTITUTIONEN

**Clinical Outcomes in Routine Evaluation-Outcome Measure
(CORE-OM) som utfallsmått i studentledd terapi**

Josefin Holmberg & Tove Regnéll

Examensarbete 30 hp
Psykologprogrammet
PM2519
Hösttermin 2019

Handledare: Kristina Berglund

Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) som utfallsmått i studentledd terapi

Josefin Holmberg & Tove Regnéll

Sammanfattning. Syftet med studien var att undersöka självskattningsinstrumentet Clinical Outcomes in Routine Evaluation-Outcome Measure:s (CORE-OM) funktion som utfallsmått för patienter som går i studentledd terapi. Deltagarna var 2223 patienter som gått i studentledd terapi samt deltagit i kvalitetssäkring vid före- och/eller eftermätning. Instrumentets validitet, reliabilitet och faktoriella struktur undersöktes. CORE-OM uppvisade god intern konsistens, hög konvergent validitet, god diskriminativ validitet förutom vad gäller specifik fobi och hög sensitivitet för förändring. Vi fann inga skillnader mellan män och kvinnors medelvärden men små negativa samband mellan medelvärde och ålder. Alla dimensioner korrelerade högt förutom Riskdimensionen. Snarare än de fyra tilltänkta dimensionerna fann vi att en trekomponentslösning förklarade datan bäst. CORE-OM bedöms vara lämpligt som utfallsmått för studentledd terapi förutom för specifik fobi-patienter.

Efter att nyttan av psykoterapi som behandling av psykisk ohälsa började att bevisas i metastudier under 1970-talet har intresset för att identifiera effektiva behandlingar och mäta dess effekter ökat (APA Presidential Task Force on Evidence-Based Practice, 2006). Evidensbaserad behandling definieras av APA som "integrationen av den bästa forskningsevidensen med klinisk expertis och patientvärden" (APA Presidential Task Force on Evidence-Based Practice, 2006, s. 273). Evidensbaserad praktik betonar vikten av att överföra forskningsresultat till klinisk vardag (Katsikis, 2014).

För att i forskningen kunna undersöka olika behandlingars effektivitet krävs instrument som kan fånga upp och kvantifiera de förändringar som behandlingen avser att åstadkomma (Barkham et al., 1998). Även klinikern är i behov av sådana instrument som kan användas till att komplettera klinisk bedömning och utvärdera behandlingsutfall (APA Presidential Task Force on Evidence-Based Practice, 2006). En typ av instrument som är vanligt förekommande både inom behandlingforsknin g och klinisk praktik är självskattningsinstrument (Barkham et al., 1998), dvs. formulär som patienten/deltagaren själv fyller i som sedan genererar en poäng på en skala. Självskattningsinstrument är mindre tidskrävande än kliniska intervjuer. Forskning har också visat att ett statistiskt instrument ofta överträffar klinikers bedömning av psykiska problems svårighetsgrad (Dawes, Faust & Meehl, 1989). Självskattningsinstruments förmåga att etablera en baslinje som kan användas som jämförelse efter en intervention gör dem användbara både i behandlingforsknin g och i klinisk vardag.

För att behandlingforsknin g och klinisk praktik ska kunna jämföras med varandra är det centralt att man använder sig av liknande utfallsmått (Barkham et al., 1998). Så var länge inte fallet då man inom forskningen tenderade att använda sig av mycket omfattande självskattningsformulär som varken var tillräckligt användarvänliga eller enkla nog att administrera för att smidigt kunna användas i klinisk praktik (Barkham et

al., 1998; 2001). Medan det i klinisk praktik fanns ett överflöd av formulär som användes utan systematik, varav flera inte var utvärderade med avseende på validitet och reliabilitet.

En forskargrupp i Storbritannien såg behovet av ett transdiagnostiskt självskattningsinstrument som både fångade in de teoretiska konstrukt som man är intresserad av att mäta i forskning och som var smidigt att använda i klinisk praktik (Barkham et al., 1998). Istället för att använda sig av olika självskattningsformulär vid olika tillstånd (t. ex. ett depressionsformulär för att mäta grad av depression) ville de konstruera ett mer generellt instrument som var användbart för att utvärdera utfall för alla sorters psykologiska problem. Detta generella instrument skulle, snarare än att utesluta behovet av andra instrument, kunna kompletteras med mer diagnosspecifika sådana (Barkham et al., 1998). De tyckte även att de var viktigt att ett sådant instrument skulle vara kostnadsfritt. Därför utvecklade de Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM, hädanefter CORE) som presenterades för första gången 1998 (Evans et al., 2000). CORE är ett transdiagnostiskt mätinstrument som är utvecklat för att fånga upp en individs mående och övervaka psykologisk förändring över tid (Barkham et al., 2001; Barkham, Clark & Stiles, 2015). Det består av 34 item som är tänkt att mäta fyra dimensioner: Subjektivt välbefinnande, Problem/Symtom, Fungerande samt Risk för andra och sig själv (se bilaga 1).

CORE har fått stort genomslag i Storbritannien (Barkham et al., 2001). Det har även översatts till 52 språk (Zeldovich & Alexandrowicz, 2019), däribland svenska, och används inom primärvård och psykiatri i Sverige. CORE används också vid studentledd psykoterapi vid universiteten i Göteborg, Lund, Umeå och Växjö. Instrumentets psykometriska egenskaper har undersökts i icke-kliniska populationer samt inom primärvård och psykiatri och genomgående uppvisat god validitet och reliabilitet (se bl. a. Evans et al., 2002; Connell et al., 2007 samt Elfström et al., 2013). Dock har CORE:s validitet och reliabilitet ej undersökts i den patientgrupp som går i studentledd terapi vid psykolog- och psykoterapeutprogrammen, varken i Sverige eller utomlands.

Validering av CORE

Kliniska gränsvärden. Engelska normer uppmättes senast av Connell et al. (2007) där man undersökte uppmätta värden i en allmän befolkning i England och sedan jämförde dessa med tidigare insamlade kliniska data (Barkham et al., 2001; Barkham, Culverwell, Spindler & Twigg, 2005; Evans et al., 2002). Det kliniska gränsvärdet på CORE beräknades till 1,0 i medelvärde (Connell et al., 2007) medan man i det svenska urvalet fann ett gränsvärde på 1,2 för män och 1,26 för kvinnor (Elfström et al., 2013). Detta menade Elfström et al. bör ses på med försiktighet då det icke-kliniska urvalet bestod av studenter vilka sannolikt skattar något högre än normalpopulationen, delvis som en avspegling av att studenter mår sämre men också för att de är mer ”psykologiskt medvetna” än normalpopulationen (Elfström et al., 2013). Elfström et al. framhöll att ett något lägre gränsvärde sannolikt hade uppmätts ifall urvalet hade varit helt slumpmässigt insamlat och att kliniker därför bör överväga både det svenska gränsvärdet och det brittiska uppmätt av Connell et al. vid bedömning. Då riskdimensionen skiljer ut sig från övriga dimensioner på flera sätt, bland annat då den uppvisar lägst korrelation med övriga dimensioner, har det rekommenderats att man också överväger totalpoängen exklusive

Risk vid bedömning (Lyne, Barrett, Evans & Barkham, 2006). När Elfström et al. (2013) utslöt Risk fann de gränsvärden på 1,41 för män och 1,49 för kvinnor i medelvärde.

Validitet och reliabilitet. CORE har undersökts i syfte att se om det som mätinstrument uppvisar reliabilitet (intern konsistens och test-reteststabilitet), validitet (konvergent validitet samt skillnader mellan kliniska och icke-kliniska populationer), acceptabilitet (dvs. om det är användarvänligt och få item missas) och sensitivitet för förändring (huruvida instrumentet fångar upp förändringar i en individs mående) (Zeldovich & Alexandrowicz, 2019; se också Evans et al., 2002). Evans et al. (2002) gjorde den första studien där man utvärderade instrumentet på detta sätt och de kunde konkludera att CORE uppvisar reliabilitet, validitet och känslighet för förändring; att det är acceptabelt och tillgängligt samt att det har en bred klinisk användbarhet. Denna studie har sedan använts som mall i flera andra studier när man har översatt och utvärderat instrumentet i olika länder (t. ex. Campbell & Young, 2011; Elfström et al., 2013; Kristjánsdóttir et al., 2015; Trujillo et al., 2016; Viliūnienė et al., 2013).

Zeldovich och Alexandrowicz (2019) har i en metastudie samlat resultaten från 19 olika översättningar av CORE, däribland Elfström et al. (2013). I studien har man bland annat tittat på reliabilitet (intern konsistens mätt med Cronbachs alfa och test-reteststabilitet) och validitet (konvergent validitet och faktoranalys). Zeldovich och Alexandrowicz fann en genomsnittlig alfakoefficient som spände mellan 0,72 (för Subjektivt välbefinnande) och 0,93 (på samtliga item). Subjektivt välbefinnande och Risk uppvisade lägre intern konsistens än övriga dimensioner (Zeldovich & Alexandrowicz, 2019). Connell et al. (2007) som inte är inkluderade i Zeldovich och Alexandrowicz metastudie fann i sitt stickprov från allmänbefolkningen även de en hög intern konsistens på totalpoängen ($\alpha=0,91$).

Man har funnit att CORE uppvisar god test-reteststabilitet (Zeldovich & Alexandrowicz, 2019; också Barkham, Mullin, Leach, Stiles & Lucock, 2007; Evans et al., 2002). Vidare har konvergent validitet beräknats med hjälp av Beck Depression Inventory-II (BDI-II) där man såg att den genomsnittliga korrelationen var $r=0,84$ (Zeldovich & Alexandrowicz, 2019). Lägst korrelation fann man mellan BDI-II och Riskdimensionen där genomsnittet var $r=0,53$ (Zeldovich & Alexandrowicz, 2019). Man har inte jämfört den svenska versionen av CORE med BDI-II men Elfström et al. (2013) undersökte den svenska versionens konvergenta validitet med hjälp av Hospital Anxiety Depression Scale (HADS) (Elfström et al., 2013). Man fann att hela CORE korrelerade med 0,66 med HAD depression och 0,68 med HAD ångest (Elfström et al., 2013).

Samband mellan CORE-poäng, kön och ålder. Man har i flera studier sett mycket små, men signifikanta negativa samband mellan poäng på CORE och ålder (Evans et al., 2002; Connell et al., 2007; Elfström et al., 2013; Trujillo et al., 2016). I Sverige har man framförallt sett tendenser till att poängen på Risk och Fungerande minskar med ålder (Elfström et al., 2013). Det har diskuterats huruvida ett lägre cut-off-värde bör användas på äldre individer, då en klinisk grupp bestående av deltagare som var 65 år eller äldre uppvisade betydligt lägre normer än de man hittat i tidigare studier (Barkham, et al., 2005). Vissa studier har dock visat upp ett positivt samband mellan ålder och poäng i en icke-klinisk population (Evans et al., 2002; Barkham et al., 2005) medan man i andra inte funnit några samband med ålder alls (Kristjánsdóttir et al., 2015; Palmieri et al., 2009).

De könsskillnader man har sett som är återkommande är högre skattningar på Risk hos män (Elfström et al., 2013; Skre et al., 2013 NY; Uji, Sakamoto, Adachi & Kitamura, 2012) och högre skattningar på Subjektivt välbefinnande (dvs. sämre välbefinnande) hos

kvinnor (Evans et al., 2002; Trujillo et al., 2016; Uji et al., 2012). I Sverige har man kommit fram till olika gränsvärden för män och kvinnor (Elfström et al., 2013) medan man i andra länder kommit fram till ett gemensamt gränsvärde för båda könen (Connell et al., 2007; Kristjánsdóttir et al., 2015).

Faktorstruktur. Gällande CORE:s faktoriella struktur har man med faktoranalys inte kunnat hitta de fyra dimensioner som CORE är tänkt att mäta (Bedford et al., 2010; Lyne et al., 2006; Zeldovich & Alexandrowicz, 2019). I metaanalysen gjord av Zeldovich och Alexandrowicz var det ingen av de åtta studier som utförde en principalkomponentanalys (PCA) som fann en fyrfaktoriell lösning tillfredställande. Däremot föreslog fyra av studierna likt Evans (2002) en indelning i tre dimensioner i form av positivt formulerade item, negativt formulerade item och item från Riskdimensionen (Zeldovich & Alexandrowicz, 2019). Bedford et al. (2010) prövade också att extrahera fyra komponenter vid en PCA vilket ej genererade de fyra hypotetiska dimensionerna. Inte heller när man genomfört en konfirmatorisk faktoranalys för att pröva hypotesen att CORE utgörs av de fyra tilltänkta dimensionerna fann någon studie att den strukturen var den bästa grupperingen (Zeldovich & Alexandrowicz, 2019).

Lyne et al. (2006) påpekade att det är svårt att skilja ut dimensionerna Subjektivt välbefinnande, Problem/Symtom och Fungerande ifrån varandra då de korrelerar starkt och alla har en stark koppling till den mer generella faktorn psykisk ohälsa. I linje med detta fann Bedford et al. (2010) i sin PCA att resultatet blev en komponent bestående av ett stort antal item av psykisk ohälsa och en mindre andra komponent som bestod av item från Risk. När de utförde en PCA på CORE:s totalpoäng exklusive Risk utgjordes istället den andra komponenten av item från Fungerande. Detta menade Bedford et al. talar för att inte bara Risk utan även Fungerande bör separeras från resten av CORE då både Risk och Fungerande skiljer ut sig från den generella komponenten psykisk ohälsa. De konkluderade till och med att det var teoretiskt felaktigt att kombinera item som mäter symtom/problem med relaterande och fysiskt våld i ett enda formulär eftersom de tillhör olika "diskursuniversum" (Bedford et al., 2010, s. 60). Eftersom den första komponenten (psykisk ohälsa) i båda fallen av PCA (med och utan Risk) innehöll så många item framhöll de även att det lämpligaste vore att minska antalet item som mäter psykisk ohälsa då de i för hög utsträckning mäter samma sak.

Riskdimensionen. Riskdimensionen har genomgående skiljt ut sig från de övriga dimensionerna och från skalan som helhet (Zeldovich & Alexandrowicz, 2019). Riskdimensionen är den dimension som korrelerar lägst med övriga dimensioner och med hela skalan samt uppvisar lägre test-retest-reliabilitet och lägre korrelation med instrument som använts för att uppmäta konvergent validitet, däribland BDI-II. Utifrån sin metastudie konkluderade Zeldovich och Alexandrowicz att Risk uppvisar stora problem i den kliniska population, dvs. primärvårdskontext, som CORE är tänkt att användas inom, vilket kan illustreras av att Elfström et al. (2013) noterade att enbart ett fåtal deltagare i deras studie svarade positivt på vissa riskitem. Däremot föreslog Zeldovich och Alexandrowicz att riskdimensionen kan tänkas fungera bättre i en psykiatrisk kontext där patienterna är inneliggande och därmed kan tänkas utgöra mer av en Risk för sig själv och andra. Kristjánsdóttir et al. (2015) påpekade att ett problem med riskdimensionen kan vara att den avser att fånga in två konstrukt (risk mot andra och risk för den egna personen) vilka inte nödvändigtvis är nära besläktade.

Även resultat från faktoranalys har, som nämnts ovan, visat att riskdimensionen utmärker sig jämfört med de övriga dimensionerna och det har föreslagits att CORE kanske bäst bör conceptualiseras som en riskdimension och en dimension för generell

psykisk ohälsa (Bedford et al., 2010; Lyne et al., 2006; Zeldovich & Alexandrowicz, 2019). Evans et al. (2002) föreslog att man kan utvärdera totalpoängen exklusive riskdimensionen och Zeldovich och Alexandrowicz (2019) instämde i detta. Lyne et al. (2006) framhöll att det mest användbara sättet att beräkna poängen är med Risk för sig och resten av totalen (28 item) för sig, och att CORE uppvisar god kvalitet ifall man använder den på det sättet.

Syfte och frågeställningar

Sammanfattningsvis har CORE visat sig ha gott stöd för att användas som utfallsmått för psykoterapi, även om det har framförts kritik gällande instrumentets faktorstruktur och riskdimension. Studier där man undersökt CORE:s validitet och reliabilitet har gjorts i flera länder med olika urval men aldrig på den grupp patienter som går i studentledd terapi, trots att CORE används som utfallsmått på studentmottagningar vid flera universitet i Sverige. Med avstamp i detta är syftet med denna undersökning att utvärdera CORE som utfallsmått för patienter som går i studentledd terapi. Specifikt ska vi undersöka CORE:s interna konsistens, konvergenta validitet, diskriminativa validitet, köns- och ålderskillnader, korrelationer mellan dimensioner, faktoriella struktur samt sensitivitet för förändring i denna grupp.

Metod

Deltagare

Deltagarna i studien var patienter som gick i studentledd psykoterapeutisk behandling på en universitetsmottagning mellan 2013 och 2018 och som deltog i kvalitetssäkringen vid före- och/eller eftermätning. Totalt har 2 223 personer fyllt i antingen före- eller eftermätning. Av dessa hade 1 793 personer fyllt i föremätning och 1 544 fyllt i eftermätningen. Totalt hade 1 114 personer fyllt i båda mätningarna. Av de 2 223 deltagarna besvarade 1 352 personer frågan om ålder vilken varierade mellan 18 och 94 år med ett medelvärde på 32,20 år (standardavvikelse [s]=10,34). Vad gäller kön besvarade 1 361 personer frågan och av dessa var 1 050 (77,1%) kvinnor och 311 (22,9%) män.

Av deltagarna var det 430 personer som inte angav vilken behandlingsform de gick i. Det var 1 057 deltagare som angav att de hade gått i kognitiv beteendeterapi (KBT), 360 att de hade gått korttidsdynamisk terapi (BDT) och 376 att de hade gått i psykodynamisk terapi (PDT). Enbart de deltagare som gått i KBT tillfrågades om vilket psykiskt problem de hade behandlats för (i en webbaserad ettårsuppföljning), och av dessa hade 274 svarat på frågan. Av dessa hade 17 deltagare behandlats för paniksyndrom, 23 för social ångest, 87 för specifik fobi, 24 för depression, 16 för tvångssyndrom, 60 för oro/ångest, 11 för stressproblematik, 6 för sömnproblem och 30 hade behandlats för annan problematik än ovanstående alternativ.

Instrument

CORE. CORE består av totalt 34 item som är uppdelade i fyra dimensioner och är tänkt att mäta individens upplevda mående under den senaste veckan (Evans, 2019). De fyra dimensionerna heter Subjektivt välbefinnande (hädanefter: Välbefinnande), Problem/Symtom, Fungerande och Risk. Namnen på dimensionerna avspeglar vad de är tänkta att mäta. Problem/Symtom, Fungerande och Risk delas in i underkategorier medan Välbefinnande utgör ett enda konstrukt (Barkham et al., 2001). Välbefinnande består av totalt fyra item som berör hur individen har upplevt sig må under veckan som gått. Problem/Symtom består av tolv item som är tänkta att mäta ångestproblem, depressiva problem, fysiska somatiska problem och traumarelaterade problem. Fungerande består av tolv item som är tänkta att mäta fungerade i nära relationer, allmänt och socialt. Dimensionen Risk består av sex item tänkta att mäta risker för den egna personen och risker mot andra.

Man har balanserat formuläret så att 18 item speglar hög intensitet i symtomen och 16 item speglar en lägre intensitet för att kunna fånga upp en bredd av intensitet sett till symtom och påverkan (Evans, 2019). Ett exempel på ett påstående med hög intensitet från de item som mäter depressionssymtom är: "Jag har känt förtvivlan eller hopplöshet" (Evans et al., 2002). Ett exempel från samma skala på ett item som speglar lägre intensitet i symtomen är "Jag har känt mig olycklig". Svartalternativen på varje item är "Aldrig", "Sällan", "Då och då", "Ofta" och "Nästan hela tiden" (Evans et al., 2000). Poängsättningen är 0-4 på varje fråga och högre poäng indikerar sämre mående. Åtta item är formulerade så att de behöver vändas efter skattning (Evans, 2019). Om några item har missats, dock högst tre, rekommenderar Evans et al. (2002) att man ersätter de saknade skattningarna med medelvärdet av poängen på de item som är ifyllda. Totalpoängen kan anges antingen i medelvärde eller i så kallad klinisk poäng vilket är medelvärdet multiplicerat med tio (Evans, 2019).

Beck Depression Inventory-II (BDI-II). BDI-II är ett självskattningsformulär som mäter depression hos vuxna och ungdomar (Beck, Steer & Brown, 2005). Formuläret består av 21 item som skattas från 0 till 3. Svartalternativen är olika på varje fråga men har gemensamt att intensiteten i symtomet stegrar för varje svartalternativ. På exempelvis item nummer ett som avser mäta nedstämdhet är alternativen: 0="Jag känner mig varken ledsen eller nedstämd", 1="Jag känner mig ofta ledsen och nedstämd", 2="Jag känner mig ledsen och nedstämd hela tiden" och 3="Jag är så ledsen och olycklig att jag inte står ut". BDI-II har uppvisat god intern konsistens, test-reteststabilitet, innehållsvaliditet och konvergent validitet. Endast legitimerad psykolog eller läkare eller annan med motsvarande kompetens, eller som har översyn av en person med en sådan kompetens, får lov att administrera instrumentet. BDI-II är enkel och tidseffektiv att fylla i. Maxpoängen är 63 och man har kommit fram till följande gränsvärden: 0-13 minimal depression, 13-19 lindrig depression, 20-28 måttlig depression och 29-63 svår depression. Formuläret är två A4 långt.

Beck Anxiety Inventory (BAI). BAI är ett självskattningsformulär som mäter ångestnivå hos vuxna och ungdomar (Beck & Steer, 2012). Det består av 21 item som kan besvaras 0="Inte alls", 1="Lite (Har inte besvärat mig mycket", 2="Delvis (Har varit väldigt obehagligt, men jag kunde uthärda det)" eller 3="Mycket (Har knappt stått ut)". Till skillnad från BDI-II avser inte BAI stämman överens med kriterierna för en specifik diagnos utan mäter förekomst av ångestsymtom vilka kan finnas närvarande vid många diagnoser, t. ex. paniksyndrom, generaliserat ångestsyndrom, depression och social fobi.

BAI har också uppvisat god validitet och reliabilitet i form av intern konsistens, test-reteststabilitet och innehållsvaliditet. Vad gäller konvergent validitet så korrelerar BAI signifikant med andra mått som avser att mäta ångest, dock är korrelationerna lägre än mellan BDI-II och mått som avser mäta depression (Beck, Steer & Brown, 2005; Beck & Steer, 2012). BAI:s 21 frågor får alla plats på en A4-sida och maxpoängen är även här 63. Gränsvärden är som följer: 0-7 minimal ångest, 8-17 lindrig ångest, 16-25 måttlig och 26-63 svår ångest.

Tillvägagångssätt

På psykoteraimottagningen bedriver studenter på psykolog- och psykoteraiprogrammen psykoteraipisk behandling under handledning. Terapiformerna som erbjuds är psykodynamisk terapi (PDT), kognitiv beteendeterapi (KBT) och Brief Dynamic Therapy (BDT). Terapiernas längd varierar mellan en och tre terminer.

Patienterna ansöker själva om att få gå i terapi genom den psykologiska institutionens hemsida. Den som blir kallad får först träffa en handledare eller anställd på psykoteraimottagningen för en bedömning ifall man är lämpad för studentledd terapi. Om så är fallet blir man sedan kontaktad av sin studentterapeut och behandlingen kan inledas.

Kvalitetssäkringen går till så att patienten blir tillfrågad att fylla i CORE vid det initiala bedömningstillfället samt vid terapiavslut. De fyller också i vissa bakgrundsfrågor (t.ex kön och ålder). Även studentterapeuten fyller vid terapiavslut i ett formulär med frågor om hur hen tycker att behandlingen har gått. Datainsamlingen har pågått sedan 2011, men omarbetades 2013. Sedan 2013 har även de patienter som samtyckt till att bli kontaktade efter avslutad terapi medverkat i en uppföljningsmätning cirka 12 månader efter avslutad behandling. Även då fyller de i CORE samt ett antal frågor om bland annat vilket problem de behandlades för, huruvida de fullföljde terapin samt ifall de för närvarande går i psykoteraipisk eller medicinsk behandling. Vid uppföljningsregistreringen ombeds även de som har gått i KBT att ange vad för psykiskt problem som de huvudsakligen behandlats för (social fobi, specifik fobi, depression, tvångssyndrom, oro/ångest, stress, sömnproblem eller annat). För de patienter som går i PDT/BDT registreras inte psykiskt problem över huvud taget i kvalitetssäkringen.

Nytt för hösten 2019 är att patienten efter fjärde samtalet fyller i ett antal frågor om hur behandlingen har fungerat hittills. Då fyller patienten dock inte i CORE. Inga andra instrument används specifikt till kvalitetssäkringen oavsett problemområde eller terapiform. Dock fyller alla KBT-patienter i Beck Anxiety Inventory (BAI) och Beck Depression Inventory-II (BDI-II) vid start och avslut av behandling. Resultaten på dessa instrument har emellertid inte samlats in systematiskt. Utöver BAI och BDI-II är det upp till varje studentterapeut att ge ut de formulär till sina KBT-patienter som de finner lämpliga i syfte att samla information och mäta framsteg. Detta sker dock inom ramen för behandlingen och samlas inte heller in till kvalitetssäkringen. PDT- och BDT-patienterna fyller inte regelmässigt i några självskattningsformulär utöver CORE.

Den insamlade datan används för att utvärdera terapieffekterna på kliniken och har legat till grund för flertalet studentuppsatser. Studentterapeuten får inte tillgång till sina patienters resultat på CORE utan datan används enbart till att mäta terapieffekter på gruppnivå. Patientdatan som användes till analyserna erhöles från institutionens ansvariga

för kvalitetssäkring. Förutom före- och eftermätningarna på CORE ingick i vårt dataset även ett mindre antal patienters ($N=151$) resultat på före- och eftermätningar av BDI-II och BAI som samlats in till kvalitetssäkringen.

Databearbetning och statistik

Vi gick först igenom datan för att se om det fanns något bortfall, och fann att fem deltagare från föremätningen fick uteslutas då de inte hade svarat på fler än tre item. Elva deltagare hade vid föremätningen missat att fylla i ett eller två item vilka ersattes med medelvärdet av de item som deltagaren i fråga har svarat på i enlighet med Evans et al. (2002). Vid eftermätningen var det ingen som hade missat fler än tre item, medan 24 deltagares resultat ersattes med medelvärden då de hade missat ett eller två item.

Vidare granskades resultaten från BDI-II och BAI. Optimalt för vår frågeställning hade varit om patienten ombetts fylla i de tre formulären vid samma tillfälle, men så har det inte sett ut. Istället har CORE delats ut i samband med det initiala bedömningsamtal som handledaren har innan patienten får träffa studentterapeuten som i sin tur delat ut BDI-II och BAI. Det kan gå olika lång tid mellan dessa mätningar. Vid avslut är det dock studentterapeuten som ansvarar för att dela ut både CORE samt BDI-II och BAI, så även om dessa inte administreras vid exakt samma tillfälle kan man anta att de åtminstone fyllts i närmare varandra i tid än vid föremätningen. Därför har vi valt att beräkna den konvergenta validiteten på datan från avslutsregistreringen. Efter att de 42 deltagare som inte svarade vid andra tillfället samt de 15 som svarat på BDI-II och BAI men inte på CORE sorterats bort återstod 94 deltagare i detta urval.

Analyserna har gjorts på det urval som har genererat störst N , alternativt passat bäst för analysen i fråga. Exempelvis utfördes beräkningarna för inre konsistens samt köns- och ålderskorrelationer på föremätninggruppen (p.g.a. störst N) medan den konvergenta validiteten beräknades med hjälp av eftermätningarna av CORE, BDI-II och BAI eftersom de vid avslut administrerades närmare varandra i tid än vid föremätningarna.

Testerna som användes var beräkning av alfakoefficienter för att mäta intern konsistens, t-test med ett stickprovsmedelvärde för att mäta diskriminativ validitet, t-test för oberoende mätningar för att beräkna könsskillnader, t-test för beroendemätningar för att undersöka sensitivitet för förändring, Pearsons produktmomentkorrelationskoefficient för att mäta konvergent validitet, Spearmans rangordningskorrelation för att mäta samband med ålder och korrelationer mellan dimensioner samt principalkomponentsanalys för att undersöka den underliggande faktorstrukturen. Signifikansnivån sattes genomgående till 0,05.

Resultat

Intern konsistens

Cronbachs alfa beräknades för alla item, alla item exklusive Risk samt alla dimensioner var för sig på patientskattningarna från föremätningen ($N=1\ 793$). De

fullständiga resultaten redovisas i tabell 1. Alfakoefficienten varierade mellan 0,75 och 0,94 vilket är acceptabelt.

Tabell 1. Alfakoefficienten α , som betecknar intern konsistens, för CORE:s olika dimensioner vid föremätningen ($N=1793$).

Dimension (antal item)	α
Välbefinnande (4)	0,76
Problem/Symtom (12)	0,89
Fungerande (12)	0,84
Risk (6)	0,75
Alla (34)	0,94
Alla utom Risk (28)	0,94

Konvergent validitet

För att utvärdera ifall CORE uppvisade konvergent validitet i detta urval beräknades CORE:s korrelationer med BDI-II och BAI. BDI-II och BAI är två instrument som mäter depression och ångest vilket teoretiskt överlappar med det som CORE ska mäta, dvs. generell psykisk ohälsa. För att CORE ska uppvisa konvergent validitet bör resultaten alltså visa att CORE och instrumenten i fråga korrelerar. Detta har undersökts i tidigare forskning och man har funnit att instrument som mäter depressionssymtom korrelerar högre med CORE än instrument som mäter ångest (Barkham et al., 2006; Evans et al., 2002). Då föremätningen med CORE sker vid bedömningssamtal och föremätningen med BDI och BAI sker efter första samtalet med tilldelad terapeut så kan tiden som förflyter mellan dessa två tillfällen vara olika lång. Eftermätningen med BDI-II och BAI sker i regel i samband med det avslutande samtalet och vid samma session får patienterna med sig CORE att fylla i på egen hand. Därmed är det troligt att tidsperioden från att patienterna fyller i eftermätningen av BDI-II och BAI till att de fyller i eftermätningen av CORE i genomsnitt är kortare än tiden som förflyter mellan föremätningarna på instrumenten. För att få skattningarna på instrumenten så jämförbara som möjligt valde vi därför att jämföra eftermätningarna och undersöka korrelationerna mellan CORE, BDI och BAI baserat på dessa ($N=94$).

Vi fann höga korrelationer mellan BDI-II och CORE på totalpoängen både med och utan riskdimensionen ($r=0,79$, $p<0,001$). Korrelationen till enskilda dimensioner varierade mellan 0,66-0,77, bortsett från Risk som låg på 0,32. Dessa resultat är jämförbara med Evans et al. (2002) som fann samma korrelation på Risk, 0,85 på alla item och 0,74-0,79 på övriga dimensioner. Korrelationen mellan BAI och CORE var genomgående lägre. Korrelationen för totalpoängen med och utan riskdimensionen låg på 0,53 ($p=<0,001$). Korrelationerna för enskilda dimensioner varierade mellan 0,57-0,21, varav Problem/Symtom uppvisade högst korrelation medan korrelationen med riskdimensionen var lägst. Dessa korrelationer är lägre än vad Evans et al. (2002) fann men följer i övrigt deras resultat. Att den högsta korrelationen var för problem/symtom var väntat då BAI framförallt mäter symtom (Beck & Steer, 2012).

Tabell 2. Korrelationer (Pearsons r) med BDI-II och BAI vid eftermätningen ($N=94$).

Mätinstrument	Dimension					
	V	P/S	F	R	Alla-R	Alla
BDI-II	0,71***	0,77***	0,66***	0,32**	0,79***	0,78***
BAI	0,41***	0,57***	0,41***	0,21*	0,53***	0,53***

V=Subjektivt välbefinnande; P/S=Problem/Symtom; F=Fungerande; R=Risk; Alla-R=alla item utom Risk; Alla=alla item; BDI-II=Beck Depression Inventory-II; BAI=Beck Anxiety Inventory.

* $p<0,05$.

** $p<0,01$.

*** $p<0,001$.

Diskriminativ validitet

En viktig aspekt av CORE:s lämplighet som utfallsmått för studentledd terapi är instrumentets förmåga att särskilja patienter som går i studentledd terapi från en icke-klinisk population. Då vi inte har samlat in data från en egen icke-klinisk jämförelsegrupp använde vi oss av de poäng från en icke-klinisk normgrupp ($N=229$) som uppmätts av Elfström et al. (2013) för att undersöka detta. T-test med ett stickprovsmedelvärde användes för att se ifall medelvärdet på CORE bland de patienter som går i studentledd terapi skiljde sig signifikant från medelvärdet från den icke-kliniska normgruppen. Vidare användes de standardavvikelser som Elfström et al. (2013) angett för att beräkna effektstorlekar i form av Cohens d .

Tabell 3. Skillnader mellan en icke-klinisk normgrupp ($N=229$) (Elfström et al., 2013) och ett urval patienter från studentkliniken ($N=1\ 792$), t -värden och effektstorlekar uttryckta i Cohens d .

Dimension	Icke-klinisk normgrupp		Studentkliniken		Skillnad M (95% KI)	t	d
	M	s	M	s			
V	1,15	0,75	1,87	0,82	0,72 (0,68-0,76)	37,15***	0,92
P/S	1,15	0,71	1,94	0,78	0,79 (0,76-0,83)	43,29***	1,06
F	0,94	0,60	1,39	0,61	0,45 (0,42-0,48)	31,48***	0,74
R	0,14	0,32	0,30	0,43	0,17 (0,15-0,19)	16,58***	0,42
Alla-R	1,06	0,60	1,70	0,66	0,64 (0,61-0,67)	41,12***	1,01
Alla	0,90	0,52	1,45	0,58	0,55 (0,52-0,58)	39,96***	1,00

V=Subjektivt välbefinnande; P/S=Problem/symtom; F=Fungerande; R=Risk; Alla-R=alla item utom risk; Alla=alla item; M =medelvärde, s =standardavvikelse, KI =konfidensintervall.

*** $p<0,001$.

Skillnaderna var signifikanta för alla dimensioner och redovisas i tabell 3. Enligt Cohens (1988) riktlinjer var skillnaderna stora på dimensionerna Välbefinnande, Problem/Symtom, hela skalan samt hela skalan exklusive Risk; måttlig på Fungerande och liten på Risk. På totalpoängen var effektstorleken på skillnaden 1,00, vilket innebär att 84,13% av patienterna i den kliniska gruppen skattade högre än medelvärdet i den icke-kliniska gruppen. Skillnaderna var genomgående mindre än vad Elfström et al.

(2013) fann mellan den kliniska och icke-kliniska gruppen, då patienterna på studentmottagningen skattade lägre på alla dimensioner.

För att vidare undersöka CORE:s diskriminativa validitet undersöktes i ett delurval patienter ($N=253$) CORE:s förmåga att särskilja olika patientgrupper från en icke-klinisk grupp. Vi var framför allt intresserade av specifik fobi-patienterna som det tidigare noterats skattade lägre på CORE jämfört med de andra patienterna på studentkliniken (Jarenfors & Thörne, 2018). Även här användes Elfström et al:s (2013) medelvärde från den icke-kliniska normgruppen ($N=229$) som testvärde.

Som synes i tabell 4 utmärkte sig specifik fobi-gruppen tydligt då den uppvisade ett medelvärde på 0,84 vid föremätningen, till skillnad från övriga grupper som låg mellan 1,32 och 1,68. Specifik fobi var också den enda gruppen som inte signifikant skiljde sig från den icke-kliniska normgruppen, $t(84) = -1,12$, *ns*.

Tabell 4. Skillnader i medelvärde på alla item på CORE-OM mellan ett urval patienter från studentkliniken ($N=253$) uppdelade per diagnos och medelvärdet ($=0,9$) från en icke-klinisk normgrupp ($N=229$) (Elfström et al., 2013), t-värden och effektstorlekar uttryckta i Cohens *d*.

Diagnos/problem	Skillnad från icke-klinisk normgrupp (testvärde=0,9; $s=0,52$)		<i>t</i>	<i>d</i>
	<i>M</i>	<i>s</i>		
Paniksyndrom ($N=17$)	1,52	0,65	0,63 (0,29-0,96)	3,97** 1,05
Social fobi ($N=23$)	1,68	0,52	0,79 (0,55-1,01)	7,19*** 1,50
Specifik fobi ($N=85$)	0,84	0,48	-0,06 ((-0,16)-0,05)	-1,12 (<i>ns</i>) -0,12
Depression ($N=23$)	1,67	0,33	0,77 (0,63-0,91)	11,36*** 1,77
Tvångssyndrom ($N=15$)	1,35	0,69	0,45 (0,07-0,83)	2,55* 0,74
Oro/ångest ($N=60$)	1,61	0,53	0,71 (0,58-0,85)	10,42*** 1,35
Annat ($N=30$)	1,32	0,51	0,42 (0,23-0,61)	4,5*** 0,82

M=medelvärde, *s*=standardavvikelse, *KI*=konfidensintervall, *ns*=ej statistiskt signifikant ($p>0,05$).

* $p<0,05$.

** $p<0,01$.

*** $p<0,001$.

Köns- och åldersskillnader

Samband mellan poäng på CORE vid föremätningen och kön undersöktes och som redovisas i tabell 5 fann vi ingen signifikant skillnad på totalpoängen ($N=1303$). På totalpoängen exklusive Risk var kvinnors skattningar signifikant högre, $t(1301) = -2,04$, $p<0,05$. Även på dimensionerna Välbefinnande och Problem/Symtom skattade kvinnor högre, $t(1301)=-3,65$, $p<0,01$ respektive $t(1301)=-2,82$, $p<0,01$. Även om skillnaderna i poäng går i samma riktning som de Elfström et al. (2013) fann så var det inte samma skillnader som var statistiskt signifikanta. När en Bonferroni-korrigerad för multipla t-test tillämpades blev dock inga av skillnaderna signifikanta.

Samband mellan ålder och poäng uppmättes med Spearmans rangkorrelation. Vi fann små men signifikanta negativa samband mellan ålder och poäng, både sett till totalpoäng samt för alla dimensioner separat med den högsta korrelationen för totalpoängen ($\rho=-0,14$, $p<0,001$, $N=1\ 294$) och den lägsta för Risk ($\rho=-0,11$, $p<0,001$, $N=1\ 294$).

Tabell 5. Skillnader i CORE-poäng mellan män ($N=299$) och kvinnor ($N=1\ 004$), t -värden och effektstorlekar uttryckta i Cohen's d .

Dimension	Män ($N=299$)		Kvinnor ($N=1\ 004$)		Skillnad		t	d
	M	s	M	s	M	95% KI		
V	1,73	0,80	1,93	0,81	-0,20	(-0,30)-(-0,09)	-3,65*	0,25
P/S	1,85	0,75	1,99	0,78	-0,14	(-0,24)-(-0,04)	-2,82**	0,18
F	0,94	0,60	1,39	0,61	0,00	(-0,07)-0,08	0,11 (<i>ns</i>)	0
R	0,14	0,32	0,30	0,43	0,01	(-0,05)-0,06	0,19 (<i>ns</i>)	0,02
Alla-R	1,64	0,63	1,72	0,66	-0,09	(-0,17)-0,00	-2,04*	0,12
Alla	1,40	0,55	1,47	0,59	-0,07	(-0,15)-0,00	-1,86 (<i>ns</i>)	0,12

V=Subjektivt välbefinnande; P/S=Problem/symtom; F=Fungerande; R=Risk; Alla-R=alla item utom risk; Alla=alla item; M =medelvärde, s =standardavvikelse, KI =konfidensintervall, *ns*=ej statistiskt signifikant ($p>0,05$).

* $p<0,05$.

** $p<0,01$.

Korrelationer mellan dimensionspoäng

Tabell 6 redogör för korrelationerna (Spearman's rangkorrelation) mellan de olika dimensionerna på CORE vid föremätningen. Som väntat uppmättes höga korrelationer mellan de flesta dimensionerna medan Risk korrelerade lägre med övriga dimensioner och hela skalan ($\rho=0,51-0,67$). Korrelationerna var likartade men generellt något högre än de uppmätta av Elfström et al. (2013) i den kliniska gruppen.

Tabell 6. Korrelationer (Spearman's ρ) mellan CORE-dimensioner vid föremätningen ($N=1\ 793$).

Dimension	V	P/S	F	R	Alla-R	Alla
V	1,00					
P/S	0,81	1,00				
F	0,75	0,73	1,00			
R	0,51	0,56	0,56	1,00		
Alla-R	0,89	0,94	0,90	0,59	1,00	
Alla	0,88	0,94	0,90	0,67	0,995	1,00

V=Subjektivt välbefinnande; P/S=Problem/Symtom; F=Fungerande; R=Risk; Alla-R=alla item utom Risk; Alla=alla item. Alla korrelationer var signifikanta på $p<0,001$ -nivå.

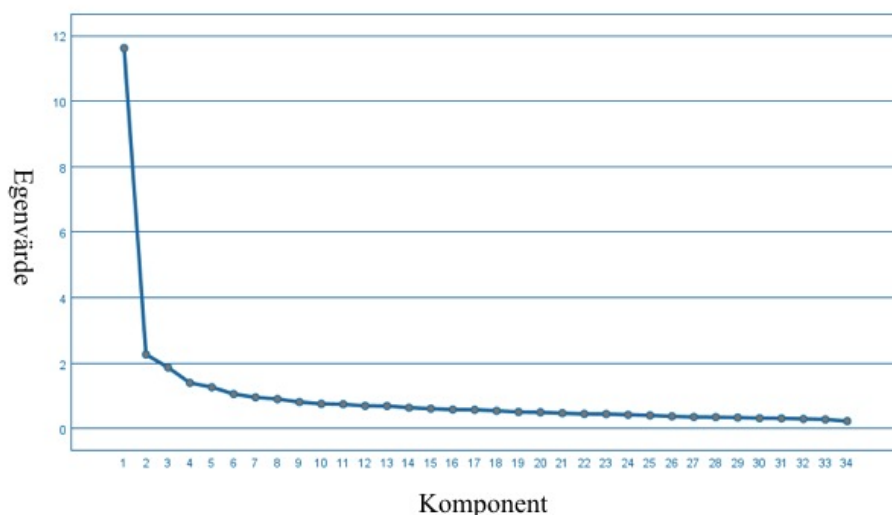
Principalkomponentanalys

För att undersöka den underliggande faktorstrukturen i CORE i detta urval genomfördes en principalkomponentanalys (PCA) på datan från föremätningen ($N=1\ 793$). Om det är så som upphovsmakarna föreslog, att CORE består av de fyra tilltänkta dimensionerna Välbefinnande, Problem/Symtom, Fungerande och Risk borde en PCA resultera i en fyrkomponentslösning där varje item laddar starkast på den komponent (dvs. dimension) som den är tänkt att tillhöra. Tidigare studier har kommit fram till olika lösningar, som nämnt ovan är en trekomponentslösning vanlig (Evans et al., 2002; Zeldovich & Alexandrowicz, 2019).

Vi valde en oblik (oblimin) rotation eftersom dimensioner i psykologiska instrument tenderar att korrelera i viss utsträckning (vilket den oblika rotationen

förutsätter) istället för den ortogonala som utgår från att dimensionerna inte korrelerar med varandra (Pett, Lackey & Sullivan, 2007). Eftersom vi gjorde en oblik rotation redogörs här för både mönster- och strukturmatrisen (se tabell 7 och 8). I mönstermatrisen redovisas varje items laddning på komponenterna. Laddning anger i en oblik rotation hur mycket en enskild komponent kan förklara variansen i ett visst item (dvs. kontrollerat för korrelationen med de övriga komponenterna). I strukturmatrisen anges itemens korrelation med komponenterna, vilka inte påverkas av korrelationen komponenterna emellan då man inte kontrollerar för dessa. Både mönster- och strukturmatrisen bör övervägas vid tolkning av en PCA med oblik rotation. Det rekommenderas att man initialt utgår från strukturmatrisen när man tolkar komponenterna, för att sedan jämföra denna med de resultat man fått i mönstermatrisen (Pett, Lack & Sullivan, 2007), vilket vi gjorde. Redovisningen av resultaten nedan utgår från laddningarna i mönstermatrisen, medan korrelationerna mellan komponenterna och itemen återfinns i strukturmatrisen (tabell 8).

Figur 1. Scree-plot.



För att avgöra hur många komponenter som var lämpligast att extrahera använde vi oss av vår scree-plot, Kaisers kriterium för egenvärde och parallellanalys. Först undersöktes scree-ploten (se figur 1) vilken kunde tolkas som att antingen tre eller fem komponenter var lämpligast. Vi använde oss av Catells kriterium som innebär att man med en linjal ritat en rak linje genom komponenterna med lägst egenvärde (Pett, Lack & Sullivan, 2007). De komponenter som ligger ovanför denna linje är de komponenter som ska extraheras. Man kan se att komponent ett står för en stor del av den totala förklarade variansen (34,17%) och att linjen sedan formar ett “knä” efter den.

Kaisers kriterium innebär att man ska ta med de komponenter som har egenvärde över ett (Pett, Lack & Sullivan, 2007). I vår analys var det sex stycken komponenter som uppfyllde detta kriterium. Horns parallellanalys är en annan metod för att fastställa antalet komponenter att extrahera. Enligt denna metod skapar man ett slumpmässigt urval av samma storlek som sitt eget urval och jämför antalet egenvärden som genereras med antalet man själv fått fram (Horn, 1965). De komponenter som får lägre egenvärde i

parallellanalysen samt uppfyller Kaisers kriterium med egenvärde över ett extraheras. Med denna metod kom vi fram till att fem komponenter skulle extraheras.

Kaisers kriterium har en tendens att hamna lite i överkant (Pett, Lack & Sullivan, 2007). Utifrån våra resultat valde vi därför att prova att begära fem, fyra samt tre komponenter. Fyr- och femkomponentslösningarna var inte tillfredsställande då enbart två item laddade över 0,4 på den fjärde faktorn. Trekomponentslösningen var mest konceptuellt begriplig och hade flera item med höga laddningar på varje faktor. Totalt 46,2% av variansen kunde förklaras med en trekomponentslösning. I likhet med tidigare forskning (Zeldovich & Alexandrowicz, 2019) delades itemen tydligt in i negativt formulerade item, positivt formulerade item och item från riskdimensionen (se tabell 7 och 8).

Tabell 7. Mönstermatrix. Itemens laddningar på de tre komponenterna.

Dimension	Item: nr	Item: förkortat innehåll	Komponent		
			1	2	3
P/S	20	Omöjligt att sluta tänka på problem	0,76		
P/S	2	Spänd, ängslig, nervös	0,74		
V	14	Känt mig gråtfärdig	0,74		
P/S	15	Känt panikkänslor/skräckslagen	0,73		
P/S	13	Upplevt oönskade tankar/känslor	0,69		
V	17	Drunkna i mina problem	0,68		
P/S	23	Känt förtvivlan/hopplöshet	0,67		
P/S	28	Påträngande minnen/bilder	0,64		
P/S	11	Spänd/ängslig hindrat mig	0,58		
P/S	27	Känt mig olycklig	0,54		
P/S	18	Svårigheter att somna in/sova hel natt	0,53		
P/S	8	Värk/smärta/andra kroppsliga problem	0,47		
P/S	5	Saknat all energi och entusiasm	0,47		
F	29	Lättretlig tillsammans med andra	0,44		
F	1	Känt mig förfärligt ensam och isolerad	0,40		
F	25	Känt mig kritiserad av andra			
R	9	Funderat på att skada mig själv		0,69	
R	34	Skadat mig själv/riskerat min hälsa		0,66	
R	16	Haft planer på att ta mitt liv		0,65	
R	6	Har brukat fysiskt våld mot andra		0,58	
R	22	Hotat/uppträtt hotfullt		0,54	
R	24	Bättre om jag var död		0,53	
F	33	Andra har förödmjukat mig		0,46	
F	32	Lyckats med det jag velat göra (vänds)			0,73
F	12	Varit nöjd med saker jag gjort (vänds)			0,70
V	4	Känt mig nöjd med mig själv (vänds)			0,69
V	31	Optimistisk inför min framtid (vänds)			0,68
F	19	Känt värme/ömheter för någon (vänds)			0,65
F	21	Klarat av det mesta jag behövt göra (vänds)			0,59
F	3	Haft någon att vända mig till (vänds)			0,55
F	7	Känt att jag klarat av situationer (vänds)			0,54
F	26	Känt som att jag inte har några vänner			
P/S	30	Tänkt att jag har mig själv att skylla			
F	10	Prata med andra alldeles för jobbigt			

P/S=Problem/Symtom, V=Subjektivt välbefinnande, F=Fungerande, R=Risk. Laddningar <0,40 censurerade.

Komponent ett, negativt formulerade item, bestod av totalt 15 item som i huvudsak hörde till Problem/symtom (11) men också Välbefinnande (2) och Fungerande (2) (se tabell 7). Laddningarna varierade mellan 0,76 och 0,40 och itemet med högst laddning

var "Det har varit omöjligt att sluta tänka på mina problem". Denna komponent genererade en Cronbachs alfa på 0,92. Komponent två bestod av sju item där ett item tillhörde Fungerande och övriga item var från riskdimensionen (6). Laddningarna varierade mellan 0,69-0,46 och itemet med högst laddning var "Jag har funderat på att skada mig själv". Cronbachs alfa för denna komponent uppmättes till 0,74 vilket är acceptabelt. Komponent tre, positivt formulerade item, utgjordes av åtta item från dimensionerna Fungerande (6) och Subjektivt Välbefinnande (2), vilka samtliga är formulerade positivt (poängen för dessa reverseras). Laddningarna varierade mellan 0,73 och 0,54 och det item med högst laddning var "Jag har lyckats med det jag har velat göra". Cronbachs alfa för denna komponent uppmättes till 0,83, vilket är bra. Fyra av itemen laddade inte över 0,40 på någon av komponenterna (se tabell 7).

Tabell 8. Strukturmatris. Itemens korrelationer med de tre komponenterna.

Dimension	Item: nr	Item: förkortat innehåll	Komponent		
			1	2	3
P/S	23	Känt förtvivlan/hopplöshet	0,81	0,41	0,55
V	17	Drunkna i mina problem	0,78		0,48
P/S	20	Omöjligt att sluta tänka på problem	0,76		0,40
P/S	2	Spänd, ängslig, nervös	0,75		0,42
P/S	27	Känt mig olycklig	0,75	0,42	0,63
V	14	Känt mig gråtfärdig	0,74		
P/S	13	Upplevt oönskade tankar eller känslor	0,73		0,41
P/S	15	Känt panikkänslor/skräckslagen	0,69		
P/S	11	Spänd/ängslig hindrat mig	0,68		0,47
P/S	28	Påträngande minnen/bilder	0,67		
F	1	Känt mig förfärligt ensam och isolerad	0,64	0,45	0,59
P/S	5	Saknat all energi och entusiasm	0,64		0,59
P/S	18	Svårigheter att somna in/sova hel natt	0,54		
F	25	Känt mig kritiserad av andra	0,52	0,49	
F	29	Lättretlig tillsammans med andra	0,52		
P/S	8	Värk/smärta/andra kroppsliga problem	0,42		
R	9	Funderat på att skada mig själv		0,75	
R	16	Haft planer på att ta mitt liv		0,69	
R	34	Skadat mig själv/risikerat min hälsa		0,65	
R	24	Bättre om jag var död	0,47	0,65	0,43
F	33	Andra har förödmjukat mig	0,41	0,55	
R	6	Har brukat fysiskt våld mot andra		0,53	
R	22	Hotat/uppträtt hotfullt		0,50	
V	4	Känt mig nöjd med mig själv (vänds)	0,51		0,78
F	12	Varit nöjd med saker jag gjort (vänds)	0,42		0,74
F	32	Lyckats med det jag velat göra (vänds)			0,74
V	31	Optimistisk inför min framtid (vänds)	0,41		0,72
F	21	Klarat av det mesta jag behövt göra (vänds)			0,61
F	7	Känt att jag klarat av situationer (vänds)			0,60
F	3	Haft någon att vända mig till (vänds)			0,57
F	19	Känt värme/ömhet för någon (vänds)			0,55
F	26	Känt som att jag inte har några vänner		0,46	0,51
P/S	30	Har mig själv att skylla	0,47		0,51
F	10	Prata med andra varit alldeles för jobbigt	0,47		0,49

P/S=Problem/Symtom, V=Subjektivt välbefinnande, F=Fungerande, R=Risk. Korrelationer <0,40 censurerade.

Sensitivitet för förändring

För att undersöka huruvida CORE är sensitiv för förändring i gruppen patienter som går i studentledd terapi beräknade vi skillnaderna mellan CORE-skattningar före och efter behandling ($N=1\ 114$). De fullständiga resultaten redovisas i tabell 9. De genomsnittliga skillnaderna på de olika dimensionerna varierade mellan 0,21 (Risk) och 0,82 (Problem/Symtom) och var alla statistiskt signifikanta ($p<0,001$). Totalmedelvärdet var över det svenska gränsvärdet för psykisk ohälsa innan behandling men under gränsvärdet efter behandling (Elfström et al., 2013). Skillnaderna var något lägre än de som Elfström et al. (2013) uppmätte förutom på riskdimensionen.

Tabell 9. Förändring i CORE-poäng ($N=1\ 114$) före och efter psykologisk behandling (KBT, PDT eller BDT).

Dimension	Före behandling		Efter behandling		Förändring: före minus efter		
	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>	<i>M</i>	95% <i>KI</i>	<i>p</i>
V	1,82	0,79	1,11	0,77	0,72	0,67-0,77	<0,001
P/S	1,91	0,76	1,08	0,71	0,82	0,78-0,87	<0,001
F	1,35	0,60	0,84	0,56	0,50	0,47-0,54	<0,001
R	0,29	0,41	0,08	0,24	0,21	0,18-0,23	<0,001
Alla-R	1,66	0,64	0,99	0,61	0,67	0,63-0,71	<0,001
Alla	1,42	0,57	0,83	0,53	0,59	0,56-0,62	<0,001

V=Subjektivt välbefinnande; P/S=Problem/symtom; F=Fungerande; R=Risk; Alla-R=alla item utom risk; Alla=alla item; *M*=medelvärde, *s*=standardavvikelse, *KI*=konfidensintervall.

Diskussion

Syftet med denna uppsats var att utvärdera CORE som utfallsmått för patienter som går i studentledd terapi. Vi ville specifikt undersöka instrumentets interna konsistens, konvergenta validitet, diskriminativa validitet, köns- och åldersskillnader, korrelationer mellan dimensioner, faktoriella struktur samt sensitivitet för förändring i denna grupp.

I våra resultat uppvisade CORE en hög intern konsistens för samtliga dimensioner utom Risk och Välbefinnande som uppvisade lägre alfakoefficienter än övriga dimensioner. Värt att notera är den höga alfakoefficienten som uppmättes för skalan som helhet. Detta går i linje med tidigare forskning (Zeldovich & Alexandrowicz, 2019). Evans et al. (2002) uppmätte samma värde som vi för skalan som helhet ($\alpha=0,94$) och beskrev den interna konsistensen som god. Dock kan ett så högt uppmätt värde peka på att det finns överflödiga item (Tavakol & Dennick, 2011).

Gällande konvergent validitet så uppvisade CORE:s totalpoäng, både med och utan Risk, en hög korrelation med BDI-II. Liknande Zeldovich och Alexandrowicz (2019) uppvisade Riskdimensionen lägst korrelation med BDI-II. Korrelationerna mellan CORE och BAI var genomgående lägre än mellan CORE och BDI-II vilket var förväntat utifrån tidigare forskning (Evans et al., 2002; Zeldovich & Alexandrowicz, 2019). Både vad gäller BDI-II och BAI fann vi lägre korrelationer med CORE än vad tidigare studier uppmätt. Detta kan vara en konsekvens av att mätillfällena för de olika instrumenten i vår studie inte har sammanfallit utan utförts vid olika tillfällen. Våra lägre korrelationer skulle också kunna bero på att vårt urval skiljde sig åt på något okänt sätt jämfört med tidigare studerade grupper. Detta har vi dock ingen anledning att tro då vi har uppmätt liknande resultat på övriga tester. Då vi inte har någon annan svensk studie att jämföra

dessa mätningar med kan vi inte heller utesluta att det beror på språkskillnader i instrumenten.

Patienterna som gick i studentledd terapi skattade signifikant högre på CORE vid föremätningen jämfört med Elfström et al:s (2013) icke-kliniska normgrupp. Dessa skillnader uppmättes i samtliga dimensioner vilket gör att vi kan dra slutsatsen att CORE uppvisar god diskriminativ validitet i detta urval, dvs. att instrumentet kan skilja patienter som går i studentledd terapi från en icke-klinisk grupp. Skillnaderna var minst för riskdimensionen, då patienterna vid föremätningen skattade mycket lågt på Risk jämfört med övriga dimensioner. Detta går i linje med att vi fann relativt låga korrelationer mellan Risk och övriga dimensioner medan alla andra dimensioner korrelerade högt med varandra. Våra resultat är i överensstämmelse med tidigare forskning (Elfström et al., 2013; Zeldovich och Alexandrowicz, 2019). Utifrån detta kan man tänka sig att upplevelse av psykisk ohälsa inte med nödvändighet hänger ihop med ett ökat riskbeteende. Zeldovich och Alexandrowicz (2019) menade att riskdimensionen inte passar i en primärvårdskontext och argumenterade för att den är lämpligare inom psykiatri med ineliggande patienter som får skatta då de antagligen i högre utsträckning är en risk för sig själva och/eller andra. Detta kan förklara varför riskdimensionen sticker ut i instrumentet även bland patienter som går i studentledd terapi då patienterna i vårt urval skattade lägre än de primärvårdspatienter som tidigare har undersökts (Elfström et al., 2013).

När olika diagnosgruppers medelvärden jämfördes med det icke-kliniska normvärde som uppmätts av Elfström et al. (2013) fann vi att specifik fobi-patienternas genomsnitt på CORE inte skiljde sig signifikant från en icke-klinisk grupp. En möjlig förklaring till detta är att den specifika fobins påverkan på individen ofta är begränsad till de situationer där fobin aktualiseras, vilket gör att fobin inte behöver ha stora negativa konsekvenser för individens generella psykiska mående. Om patienternas skattningar före en behandling redan ligger på en icke-klinisk nivå säger en skattning efter behandling ingenting om behandlingens effekter. Detta innebär att CORE inte är ett bra utfallsmått för de patienter vars primära svårighet är specifik fobi. Det hade varit lämpligt att använda sig av ett instrument som avser att mäta just specifik fobi i denna grupp för att kunna säkerställa kvaliteten även för dessa patienter.

Vi fann små negativa signifikanta samband mellan ålder och skattningar på alla dimensioner och totalpoängen, vilket går i linje med tidigare forskning (Evans et al., 2002; Connell et al., 2007; Elfström et al., 2013; Trujillo et al., 2016). Barkham et al. (2005) har föreslagit ett lägre gränsvärde för äldre individer. Våra resultat pekar också i den riktningen men vi kan inte dra några definitiva slutsatser då vårt urval dominerades av yngre deltagare (98,4% av deltagarna var yngre än 65 år).

Vi fann skillnader i skattningar mellan könen vad gäller totalpoäng exklusive Risk samt dimensionerna Välbefinnande och Problem/Symtom där kvinnors skattningar var signifikant högre. Att kvinnor skattar högre på Välbefinnande är återkommande i flera tidigare studier (Evans et al., 2002; Trujillo et al., 2016; Uji et al., 2012). De små effektstorlekarna indikerar att mäns och kvinnors resultat överlappade varandra i hög utsträckning. Skillnaderna var så små att inga av dem förblev signifikanta efter en Bonferroni-korrektion. Vi fann inte heller att män skattade högre på Risk vilket tidigare forskning har rapporterat (Elfström et al., 2013; Skre et al., 2013 NY; Uji, Sakamoto, Adachi & Kitamura, 2012). Detta pekar på att det inte finns några könsskillnader vad gäller Risk i gruppen patienter som går i studentledd terapi. Det skulle kunna bero på att målgruppen som studentkliniken riktar sig till är personer med måttlig psykisk ohälsa och

man försöker medvetet att sälla bort patienter som är suicidala, lider av missbruk eller uppvisar andra tecken på allvarlig psykisk ohälsa.

Principalkomponentanalysen resulterade i en lösning bestående av tre komponenter som uppvisade hög intern konsistens. Vi såg även att fyra item inte laddade starkt på någon komponent, vilket innebär att denna lösning bestod av 30 item totalt. De tre komponenterna bestod av negativt formulerade item, riskitem och positivt formulerade item. Detta går i linje med vad man har kommit fram till i flera andra faktoranalyser (Evans, 2002; Zeldovich & Alexandrowicz, 2019), men även en- och tvåkomponentslösningar har diskuterats (Bedford et al., 2010). Den andra komponenten utgörs som sagt huvudsakligen av riskitem, medan komponent ett och tre tycks ha att göra med huruvida itemen är vända eller inte. Det skulle kunna vara så att komponent ett är en så kallad p-faktor, dvs. generell psykopatologi (Caspi et al., 2014). Komponent tre mäter i sin tur styrkor eller upplevelser av att vara nöjd med sig själv, sina relationer och sin förmåga. Men det går inte att utesluta att de gemensamma nämnarna i båda komponenterna är hur dess item är formulerade snarare än att de mäter en viss aspekt av psykisk ohälsa. Om så är fallet är inte heller en trekomponentslösning helt tillfredsställande ifall det man är ute efter är att fånga in centrala delar i upplevelsen av psykisk ohälsa. Att det just är Risk av de tilltänkta dimensionerna som framträdde som en egen komponent tydliggör hur mycket dessa item skiljer sig från resten av instrumentet.

CORE har uppvisat sensitivitet för förändring i flertalet studier (Elfström et al., 2013; Zeldovich & Alexandrowicz, 2019) och gjorde så även i vår undersökning. Störst skillnad fann vi på Problem/Symtom och minst skillnad var det mellan före- och eftermätningarna på Risk. Som vi nämnt tidigare så har denna dimensions lämplighet ifrågasatts då skattningarna ofta varit genomgående låga. Utifrån våra resultat verkar riskdimensionen uppvisa låg klinisk användbarhet i gruppen patienter som går i studentledd terapi.

Vår undersökning har tillfört kunskap om CORE:s validitet, reliabilitet och faktoriella struktur i gruppen patienter som går i studentledd terapi, vilket inte har utforskats tidigare, varken i Sverige eller utomlands. Resultaten pekar på att CORE:s validitet, reliabilitet och faktoriella struktur på en övergripande nivå är jämförbara med vad man funnit i andra kliniska grupper i Sverige och i andra länder (Evans et al., 2002; Elfström et al., 2013; Zeldovich & Alexandrowicz, 2019). Detta trots att gruppen patienter som går i studentledd terapi verkade skilja sig något från urvalen patienter från primärvård, studentrådgivning och psykiatri som tidigare undersökts. Patienterna i vårt urval skattade generellt lägre på CORE både före och efter behandling, uppvisade lägre skillnader mellan före- och eftermätning och skiljde sig mindre från en icke-klinisk normgrupp än vad Elfström et al. och Evans et al. fann i sina urval.

Vi kunde även visa på att specifik fobi-gruppen inte skiljde sig från en icke-klinisk normgrupp, medan skillnaderna var signifikanta och stora för alla andra grupper förutom tvångssyndrom där skillnaden var medelstor enligt Cohens (1988) riktlinjer. Att denna grupp skattade lägre på CORE än övriga patientgrupper på studentmottagningen har noterats tidigare av Jarenfors & Thörne (2018) men det har inte tidigare kunnat konstateras att patienter som behandlas för specifik fobi inte skiljer sig från en icke-klinisk grupp.

Elfström et al. (2013) efterfrågade ytterligare forskning på den svenska versionen av CORE, bland annat framhöll de behovet av ytterligare undersökning av den konvergenta validiteten i ett svenskt urval då de enbart undersökt hur CORE korrelerade

med Hospital Anxiety Depression Scale (HADS) i ett mindre urval. Vi undersökte den konvergenta validiteten genom att jämföra CORE med BDI-II och BAI vilket inte tidigare har gjorts i Sverige, och fann som ovan nämnt korrelationer som var jämförbara men något lägre jämfört med tidigare forskning.

Principalkomponentanalys är ytterligare något som inte tidigare genomförts i denna grupp eller i Sverige, även om det har gjorts i andra urval på flera håll i Europa (Zeldovich & Alexandrowicz, 2019). Våra resultat liknar vad man kommit fram till tidigare (Evans et al., 2002; Zeldovich & Alexandrowicz, 2019). Vi kan dra slutsatsen att CORE:s faktoriella struktur i urvalet patienter som går i studentledd terapi ser ut att vara jämförbar med vad man funnit tidigare i andra kliniska grupper.

En möjlig brist är som tidigare nämnt att mätningarna av CORE, BDI-II och BAI inte har skett systematiskt utan sannolikt har de flesta patienterna fyllt i dessa instrument vid olika tillfällen. Detta kan förklara varför vi fann något lägre korrelationer mellan instrumenten jämfört med Evans et al. (2002).

En annan brist med vår undersökning är storleken på urvalet av patienter som angett vilken diagnos eller problematik som de gått i behandling för. Totalt hade 253 deltagare svarat på denna fråga, vilket gjorde att vissa grupper blev ganska små, t. ex. hade bara 15 deltagare angett tvångssyndrom som diagnos/problematik. Även om dessa deltagares resultat på CORE skilde sig signifikant från den icke-kliniska normgruppens är det möjligt att dessa 15 personer inte är representativa för hela gruppen patienter som går i studentledd terapi för tvångssyndrom. Detta problem har blivit åtgärdat i de uppdaterade rutinerna för kvalitetssäkring, där det istället är terapeuten som anger vilken diagnos eller problematik som varit fokus för behandlingen, vilket möjliggör en ännu noggrannare undersökning av de olika diagnosgruppernas resultat på CORE i framtiden.

Ett annat problem med urvalet var att det dominerades av kvinnor och av unga personer. För att säkerställa att våra resultat är giltiga även bland män och bland äldre personer hade det varit önskvärt med ett mer heterogent urval.

Det hade varit intressant om man i framtida forskning i ett större urval tittade närmare på CORE:s diskriminativa validitet i specifika diagnosgrupper för patienter som går i studentledd terapi. Vidare behövs de svenska kliniska gränsvärdena undersökas närmare eftersom Elfström et al:s (2013) icke-kliniska grupp utgjordes av studenter. Det vore önskvärt med ett icke-kliniskt urval som är mer representativt för Sveriges befolkning för att slutgiltigt fastställa normer och gränsvärden.

Referenser

- APA Presidential Task Force on Evidence-Based Practice. (2006) Evidence-based practice in psychology. *American Psychologist*, 61, 271-285. doi:10.1037/0003-066X.61.4.271
- Barkham, M., Culverwell, A., Spindler, K., & Twigg, E. (2005). The CORE-OM in an older adult population: Psychometric status, acceptability, and feasibility. *Aging & Mental Health*, 9, 235-245. doi:10.1080/13607860500090052
- Barkham, M., Evans, C., Margison, F., McGrath, G., Mellor-Clark, J., Milne, D., & Connell, J. (1998). The rationale for developing and implementing core outcome batteries for routine use in service settings and psychotherapy outcome research. *Journal of Mental Health*, 7, 35-47. doi:10.1080/09638239818328

- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J. & Audin, K. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Towards practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, 69, 184–196. doi:10.1037//0022-006X.69.2.184
- Barkham, M., Mellor-Clark, J., & Stiles, W.B. (2015). A CORE Approach to progress monitoring and feedback: Enhancing evidence and improving practice. *Psychotherapy*, 52, 402-411. doi:10.1037/pst0000030
- Barkham, M., Mullin, T., Leach, C., Stiles, W.B., & Lucock, M. (2007). Stability of the CORE-OM and the BDI-I prior to therapy: Evidence from routine practice. *Psychology and Psychotherapy: Theory, Research and Practice*, 80, 269-278. doi:10.1348/147608306X148048
- Beck, A.T., Steer, R.A., & Brown, G.K. (2006). *BDI-II: Beck Depression Inventory: manual, svensk version*. (2:a utg.) Stockholm: Psykologiförlaget.
- Beck, A.T., & Steer, R.A. (2012). *BAI - Beck Anxiety Inventory: manual, svensk version*. Stockholm: Pearson Sweden.
- Bedford, A., Watson, R., Lyne, J., Tibbles, J., Davies, F., & Deary, I.J. (2010). Mokken scaling and principal components analyses of the CORE-OM in a large clinical sample. *Clinical Psychology and Psychotherapy*, 17, 51-62. Hämtad 2019-11-15 från <https://search-proquest-com.ezproxy.ub.gu.se/docview/742966349?accountid=11162>
- Campbell, M.M., & Young, C. (2011). Introducing the CORE-OM in a South African context: Validation of the CORE-OM using a South African student population sample. *South African Journal of Psychology*, 41, 488-502. doi:10.1177/008124631104100408
- Caspi, A., Houts, R.M., Belsky, D.W., Goldman-Mellor, S.J., Harrington, H., Israel, S., Meier, M.H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T.E. (2014). The p Factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science*, 2, 119-137. doi:10.1177/2167702613497473
- Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2:a utg.). Hillsdale: L. Erlbaum Associates.
- Connell, J., Barkham, M., Stiles, W.B., Twigg, E., Singleton, N., Evans, O., & Miles, J.N.V. (2007). Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R. *British Journal of Psychiatry*, 190, 69-74. doi:10.1192/bjp.bp.105.017657
- Dawes, R.M., Faust, D., & Meehl, P.E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668-1674. doi:10.1126/science.2648573
- Elfström, M.L., Evans, C., Lundgren, J., Johansson, B., Hakeberg, M., & Carlsson, S.G. (2013). Validation of the Swedish version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology and Psychotherapy*, 20, 447-455. doi:10.1002/cpp.1788
- Evans, C. (2019-08-17). *Information about the CORE-OM*. Hämtad 2019-11-15 från CORE Systems Trust, <https://www.coresystemtrust.org.uk/instruments/core-om-information/>
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardized brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180, 51-60.

- Hämtad 2019-11-15 från <https://search-proquest-com.ezproxy.ub.gu.se/docview/57854825?accountid=11162>
- Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health*, 9, 247-255. Hämtad 2019-11-15 från <https://search-proquest-com.ezproxy.ub.gu.se/docview/57849382?accountid=11162>
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185. doi:10.1007/BF02289447
- Jarenfors, K., & Thörne, M. (2018). *Diagnos och behandlingsutfall: En effectiveness-studie på studentledd kognitiv beteendeterapi*. Opublicerat psykologexamensarbete, Göteborgs universitet: Psykologiska institutionen, Göteborg.
- Katsikis, D. (2014). Evidence-based research: The importance for the present and future of evidence-based practice. *Journal of Evidence-Based Psychotherapies*, 14, 259-270. Hämtad 2019-11-15 från <https://search-proquest-com.ezproxy.ub.gu.se/docview/1640009895?accountid=11162>
- Kristjánsdóttir, H., Sigurðsson, B.H., Salkovskis, P., Ólason, D., Sigurdsson, E., Evans, C., Gylfadóttir, E.D., & Sigurðsson, J.F. (2015). Evaluation of the psychometric properties of the Icelandic version of the Clinical Outcomes in Routine Evaluation-Outcome Measure, its transdiagnostic utility and cross-cultural validation. *Clinical Psychology and Psychotherapy*, 22, 64-74. doi:10.1002/cpp.1874
- Lyne, J.K., Barrett, P., Evans, C., & Barkham, M. (2006). Dimensions of variation on the CORE-OM. *British Journal of Clinical Psychology*, 45, 185-203. doi:10.1348/014466505X39106
- Palmieri, G., Evans, C., Hansen, V., Brancaloni, G., Ferrari, S., Porcelli, P., Reitano, F., & Rigatelli, M. (2009). Validation of the Italian version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology & Psychotherapy*, 16, 444-449. doi:10.1002/cpp.646
- Pett, M.A., Lackey, N.R., & Sullivan, J.J. (2003) *Making sense of factor analysis: the use of factor analysis for instrument development in health care research*. London: SAGE.
- Skre, I., Friberg, O., Elgarøy, Evans, C., Myklebust, L.H., Lillevol, K., Sørgaard, K., & Hansen, V. (2013). The factor structure and psychometric properties of the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) in Norwegian clinical and non-clinical samples. *DMC Psychiatry*, 13. doi:10.1186/1471-244X-13-99
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55. Hämtad från <https://search-proquest-com.ezproxy.ub.gu.se/docview/898889039?accountid=11162>
- Trujillo, A., Geixas, G., Bados, A., García-Grau, E., Salla, M., Medina, J.C., Montesano, A., Soriano, J., Medeiros-Ferreira, L., Cañate, J., Corbella, S., Grau, A., Lana, F., & Evans, C. (2016). Psychometric properties of the Spanish version of the Clinical Outcomes in Routine Evaluation-Outcome Measure. *Neuropsychiatric Disease and Treatment*, 12, 1457-1466. doi:10.2147/NDT.S103079
- Uji, M., Sakamoto, A., Adachi, K., & Kitamura, T. (2012). Psychometric properties of the Japanese version of the Clinical Outcomes in Routine Evaluation-Outcome

- Measure. *Comprehensive Psychiatry*, 53, 600-608.
doi:10.1016/j.comppsy.2011.09.006
- Viliūnienė, R., Evans, C., Hilbig, J., Pakalniškienė, V., Danilevičiūtė, V., Laurinaitis, E., & Navickas, A. (2013). Translating the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM) into Lithuanian. *Nordic Journal of Psychiatry*, 67, 305-311. doi:10.3109/08039488.2012.745599
- Zeldovich, M., & Alexandrowicz, R.W. (2019). Comparing outcomes: The Clinical Outcome in Routine Evaluation from an international point of view. *International Journal of Methods in Psychiatric Research*, 28. doi:10.1002/mpr.1774

Bilaga 1

Översikt över CORE-dimensioner och item.

Dimension	Item	Nr
V	Jag har känt mig nöjd med mig själv (<i>vänds</i>)	4
V	Jag har känt mig gråtfärdig	14
V	Det har känts som om jag skulle drunkna i mina problem	17
V	Jag har känt mig optimistisk inför min egen framtid (<i>vänds</i>)	31
P/S	Jag har känt mig spänd, ängslig eller nervös	2
P/S	Jag har saknat all energi och entusiasm	5
P/S	Jag har besvärats av värk, smärta eller andra kroppsliga problem	8
P/S	Att jag varit spänd och ängslig har hindrat mig från att göra viktiga saker	11
P/S	Jag har besvärats av oönskade tankar och känslor	13
P/S	Jag har haft panikkänslor eller varit skräckslagen	15
P/S	Jag har haft svårigheter att somna in eller att sova en hel natt	18
P/S	Det har varit omöjligt att sluta tänka på mina problem	20
P/S	Jag har känt förtvivlan eller hopplöshet	23
P/S	Jag har känt mig olycklig	27
P/S	Minnen eller bilder har trängt sig på och stört mig	28
P/S	Jag har tänkt att jag har mig själv att skylla för mina problem och svårigheter	30
F	Jag har känt mig förfärligt ensam och isolerad	1
F	Jag har känt att jag haft någon att vända mig till när det behövts (<i>vänds</i>)	3
F	Jag har känt att jag kunnat klara av situationer där något gått snett (<i>vänds</i>)	7
F	Jag har varit nöjd med de saker som jag har gjort (<i>vänds</i>)	12
F	Att prata med andra har känts alldeles för jobbigt	10
F	Jag har känt värme eller ömhet för någon (<i>vänds</i>)	19
F	Jag har klarat av det mesta som jag behövt göra (<i>vänds</i>)	21
F	Jag har känt mig kritiserad av andra	25
F	Det har känts som att jag inte har några vänner	26
F	Jag har varit lättretlig tillsammans med andra	29
F	Jag har lyckats med det jag velat göra (<i>vänds</i>)	32
F	Andra människor har förödmjukat mig eller fått mig att skämmas	33
R	Jag har brukat fysiskt våld mot andra	6
R	Jag har funderat på att skada mig själv	9
R	Jag har haft planer på att ta mitt liv	16
R	Jag har hotat eller uppträtt hotfullt mot någon annan	22
R	Jag har tänkt att det hade varit bättre om jag var död	24
R	Jag har själv skadat min kropp eller allvarligt riskerat min hälsa	34

V=Subjektivt välbefinnande; P/S=Problem/Symtom; F=Fungerande; R=Risk.